Draft – November 1996 to be submitted in revised form to Remote Sens. Environ.

Validation of Kernel-Driven Semiempirical Models for Global Modeling of Bidirectional Reflectance

Baoxin Hu, Wolfgang Wanner, Xiaowen Li, and Alan H. Strahler

The semiempirical, kernel-driven Ambrals BRDF model (Wanner et al., 1995) was developed for correcting and studying view and illumination angle effects of a wide variety of land covers in remote sensing applications. This model, also scheduled for use in producing a global bidirectional reflectance distribution function and albedo data product from EOS-MODIS and MISR data, is validated in this paper by demonstrating its ability to model 27 different multiangular data sets well, representing major types of land cover. The selection of the kernels used in the model is shown to relate to land cover type, and the inversion accuracy to be good in nearly all cases: the correlation coefficient between modeled and observed reflectances is larger than 0.9 for about half of the data sets and larger than 0.75 in all but one case where the observations are irregular. The average root mean squared error of the inversions is 0.034. A new kernel modeling the sun zenith angle dependence of multiple scattering is introduced and shown to improve fits for dense vegetation. Operation of the Ambrals model is demonstrated by applying it to an ASAS image on a per-pixel basis.

INTRODUCTION

With the increasing use of coarse and medium-resolution off-nadir viewing sensors producing reflectance data for global monitoring, analysis of the BRDF (Bidirectional Reflectance Distribution Function) of each pixel is becoming more and more important. The BRDF can not only be used to compare observations obtained at different angles or standardize observations to a common geometry but also to provide surface physical parameters and the boundary condition for radiative transport in the coupled atmosphere-earth system.

At the same time, global change research introduces new requirements into the modeling of the bidirectional reflectance of vegetation. Studies of the BRDF properties of vegetation have for the most part been carried out over thematically homogeneous surfaces and for a limited variety of land covers. The BRDF models developed also usually assume homogeneity of the land cover. However, to allow frequent global coverage the satellite sensors used in global change research typically do not possess high spatial resolution. For example, the U.S. National Oceanographic and Atmospheric Administration (NOAA) Advanced Very High Resolution Radiometer (AVHRR), commonly used in land surface monitoring, has a spatial resolution of 1.1 km at nadir. The Moderate Resolution Imaging Spectroradiometer (MODIS) to be launched in mid-1998 on the EOS-AM-1 platform, which will be the primary Earth Observing System (EOS) sensor for observations of terrestrial dynamics (Running et al., 1994), is a 36-channel radiometer covering 0.415–14.235 μm in wavelength with a spatial resolution ranging from 250 m to 1 km at nadir, depending on the band. For sensors of this kind, a given pixel will frequently contain a heterogeneous mixture of bare soil and vegetation canopies, or a mixture of spatially distinct types of vegetation with different structural and optical properties. Because of the global coverage provided, a very large number of different surface types will be viewed. Therefore, it is necessary to develop BRDF models of a type that can readily be applied to a variety of inhomogeneous land covers, that is flexible enough to respond to changing scenarios, and that is up to the demands of global data processing by being very rapidly invertible.

One type of BRDF model that fulfills these re-

[1]Center for Remote Sensing and Department of Geography, Boston University

Address correspondence to Baoxin Hu, Center for Remote Sensing, Boston University, 725 Commonwealth Avenue, Boston, MA 02215, USA.

Received 1 January 1999; revised 1 January 1999.

quirements very well is the semiempirical kernel-driven type originally suggested by Roujean et al. (1992) and later developed further by Wanner et al. (1995a) in form of the Ambrals BRDF model (Wanner et al., 1997). This model type has been successfully applied to a variety of remotely sensed data sets including AVHRR data (Leroy and Roujean, 1994; Li et al., 1996; Ruiz de Lope and Lewis, 1997) to correct for surface BRDF effects, for example in the vegetation index. The Ambrals model will also be used in generating the global MODIS BRDF and albedo standard data product (Strahler et al., 1996; Wanner et al., 1997).

Previously, only limited validation of the mathematical expressions used in the Ambrals model has been published (Strahler et al., 1995; Wanner et al., 1995b; Hu et al., 1996; Wanner et al., 1997). In this paper, we now present more extensive validation of the Ambrals BRDF model using various bidirectional reflectance data collected over a wide variety of surface types. Since currently almost no BRDF data sets for heterogeneously mixed land surfaces are available, validation is restricted to homogeneous cover types even though it is expected that the models will display their strength more obviously for mixed pixels. Data for mixed scenes will be more readily available in a few years.

MODELING BIDIRECTIONAL REFLECTANCE WITH KERNEL-DRIVEN MODELS

Kernel-Driven BRDF Models

In operational processing, the physical approach to BRDF modeling is problematic since it still is neither flexible enough nor computationally simple enough to allow global rapid inversions on a regular basis. It has therefore been the strategy for the three largest planned operational BRDF and albedo products, those from MODIS, the EOS Multi-Angle Imaging Spectroradiometer (MISR) (Diner et al., 1991), and the POlarization and Directionality of the Earth's Radiation instrument (POLDER) (Deschamps et al., 1994; Leroy et al., 1997) to use linear or semilinear semiempirical BRDF models. These models retain some physical interpretation while being highly capable of adapting to many BRDF shapes and being very rapidly invertible. Their number of parameters is small, usually three. MISR will be using the semiempirical RPV BRDF model developed by Rahman et al. (1993) in a form modified by Martonchik (Engelsen et al., 1996), POLDER will use the Roujean kernel-driven model (Roujean et al., 1992), and MODIS will rely on the kernel-driven semiempirical Ambrals BRDF model (Wanner et al., 1995a, 1997) that will here be validated, and will also run the empirical modified Walthall model (Walthall et al., 1985; Nilson and Kuusk, 1989) in parallel. Ambrals stands for: Algorithm for MODIS bidirectional reflectance anisotropy of the land surface.

In the kernel-driven semiempirical approach, the BRDF is modeled as a weighted sum of a volume scattering function and a surface scattering function (called kernels), and a constant (Roujean et al., 1992). These kernels are derived from approximations to physical BRDF models, so they retain a physical meaning. In model inversion, the weight given to each kernel is determined empirically by fitting to the multiangular observations made. Thus, it is the weights of the semiempirical kernels that are retrieved, characterizing the balance between volume and geometric scattering in the possibly mixed scene viewed.

Volume and surface scattering kernels are derived from physical radiative transfer models and geometric optical models by simplifying them to the following format by reasonable approximations:

$$R(\theta_i, \theta_v, \phi; \lambda) = c_1(\lambda)k(\theta_i, \theta_v, \phi) + c_2(\lambda), \quad (1)$$

where c_1 and c_2 are constants containing physical parameters, R is the modeled value of the bidirectional reflectance of surface objects, and k is the kernel function dependent only on viewing and illumination geometry; θ_i and θ_v are illumination and viewing zenith angles, ϕ the relative azimuth, and λ the wavelength. If after approximations the kernel k still contains parameters, they are set to a typical value that may vary from one kernel to the next.

A complete kernel-driven semiempirical model has the form

$$\begin{aligned} R(\theta_i, \theta_v, \phi; \lambda) &= f_{iso}(\lambda) + f_{geo}(\lambda)k_{geo}(\theta_i, \theta_v, \phi) \\ &\quad + f_{vol}(\lambda)k_{vol}(\theta_i, \theta_v, \phi), \end{aligned} \quad (2)$$

where the quantities k_{geo} and k_{vol} are the geometric-optical surface-scattering kernel and the radiative-transfer volume-scattering kernel, and the factors f_{geo}

and f_{vol} are their respective weights. The term f_{iso} is the contribution of isotropic scattering.

The Ambrals BRDF Model

The original Roujean model used a kernel derived from a single-scattering radiative transfer theory by Ross (1981), called the Ross-thick kernel, and a geometric-optical kernel for rectangular protrusions (Roujean et al., 1992). The latter kernel was found not to fit some cover types well, especially dense forest canopies (Roujean et al., 1992). The Ambrals BRDF model makes use of an improved set of kernels. The Ross-thick kernel, derived for large values of the effective scattering leaf area index, is also used, but alternately an approximation for low effective values of the leaf area index, called the Ross-thin kernel, is available (Wanner et al., 1995a). Which of the two kernels to use for a given inversion is decided a priori, is based on previous experience, or is decided such that the kernel chosen produces the least root mean squared error (RMSE) in inversion. Similarly, the Ambrals model allows for two alternate choices for the geometric-optical kernel, derived for different types of scenes (Wanner et al., 1995a). One, called the Li-sparse kernel, is an approximation to the geometric-optical mutual shadowing model by Li and Strahler (1992) for sparse ground objects, where the BRDF is mainly governed by shadow casting. The second geometric kernel, the Li-dense kernel, is determined by the bright sunlit object faces mainly visible due to mutual shadowing effects in dense ensembles of ground objects, for example dense forests.

The Ambrals model, as it is to be used for generating the MODIS BRDF/albedo data product, thus occurs in four variants that are used to optimize the inversion. These are the Ross-thin/Li-sparse, the Ross-thick/Li-sparse, the Ross-thin/Li-dense, and the Ross-thick/Li-dense modes of the model. The Li-sparse kernel is formulated for round crowns, the Li-dense kernel has prolate crowns with a diameters ratio of 2.5. Both kernels model crowns where the mean distance from the ground to the lower edge of the crown is half of the crown height.

Error Functions Used

Like all linear models, the Ambrals BRDF model can be inverted analytically through matrix inversion (Lewis, 1995), avoiding costly numerical inversion problems. In each wave band, a set of model parameters is determined through minimization of an error function,

$$e^2 = \frac{1}{N - n_p} \sum_{j=1}^{N} \frac{\left(R_j^{obs} - R_j^{model}\right)^2}{W_j}. \qquad (3)$$

Here, N is the number of observations, n_p the number of parameters of the model, R^{obs} and R^{model} are observed and modeled reflectances, respectively, and W_j is a weighting factor that may be chosen to give different weights to different observations if desired. In the absence of qualifying knowledge, W_j is commonly specified as unity. For a relative error measure, $W_j = R_j$. An absolute error measure generally ensures best accuracy for large values of the reflectance, which dominate albedo, but does not put a strong emphasis on small reflectances that may be of particular interest for some applications. A relative measure of error does not show this latter problem, but since larger absolute deviations are allowed for large values of the reflectance, albedo derived from the BRDF is likely to possess a larger absolute error as well.

A similar problem exists when attempting to find a function for minimizing the error in several wave bands simultaneously. The weight given to the error made in each band may be determined either to be equal, to favor bands with small albedos by introducing a weight in each band proportional to the size of the albedo, or by weighting according to the proportion of radiation present in each band. In this study, the RMSE determined for a specific data set and model is given by

$$RMSE = \sqrt{\frac{1}{n_b} \sum_{i=1}^{N_{band}} \frac{e^2}{w_i}}, \qquad (4)$$

with both W_j and w_i presently set to unity, the latter determining the contribution of each band to the selection of the best-fitting kernel combination of the Ambrals model; n_b is the number of wave bands.

DATA SETS USED

As mentioned earlier, kernel-driven BRDF models have been developed for applications at the global scale. They are designed to describe the bidirectional reflectance of mixed land cover types. But before more measurements for heterogeneous scenes are available, it is necessary and useful to validate the model using field-measured data over a single land cover type. A number of such data sets are available as detailed below, covering a large variety of land cover types with variations, for example, in canopy coverage and differences in leaf area index (LAI). All of these data sets are available for several different sun zenith angles. Some studies (e.g., Engelsen et al., 1996) have also used numerical BRDF forward modeling to explore model properties, but due to possible similarities in the mathematical expressions used in the modeling of both the forward and the inverse model their use for validation is limited. Table 1 summarizes basic properties of the data sets that were used.

1) BRDF data sets by Kimes

Kimes (1983) collected a series of multiangular data sets of a plowed field, a corn field, orchard grass and a grass lawn with 0, 25, 50 and 97 percent of vegetation cover, measured near Beltsville, Maryland. In situ measurements made in Northern Africa of three land covers with low and two land covers with high vegetation coverage were reported by Kimes et al. (1985). They are annual grassland, hard wheat, steppe grass, irrigated wheat, and soybean. Their corresponding coverage is 4, 11, 5, 70 and 90 percent, respectively. Helicopter measurements were performed on two kinds of high coverage forest (70 and 79 percent) dominated by pine trees in one case and hardwood trees in the other, both situated in Virginia (Kimes et al., 1986). Data collected in the red (580–680 nm) and the near-infrared (730–1100 nm) bands were used in this work, available for either three or four sun zenith angles depending on the case. The view zenith angles range from 0° to 75° in increments of 15°. Relative azimuth angles varied from 0° to 345° in increments of 45°.

2) Soybean data by Ranson

Three bidirectional reflectance data sets were collected by Ranson et al. (1985) over a commercial soybean field in West Lafayette, Indiana, on three dates during the summer of 1980 with an Exotech model 100 radiometer in four spectral bands (500–600, 600–700, 700–800, and 800–1100 nm). The view zenith angles observed at were 0°, 7°, 22°, 30°, 45° and 60°. View azimuth angles ranged from 0° to 315° in steps of 45°. On the three dates (July 18, July 25, and August 27), Vegetation coverage was 72, 83 and 99 percent.

3) Boreal forest data by Deering

Deering et al. (1995) measured the bidirectional reflectance of Old Jack Pine, Old Black Spruce and Aspen at the BOREAS forest in Canada on dates May 31, June 7 and July 21. The instrument used was PARABOLA, which allows to acquire radiance data for nearly the complete sky- and ground-looking hemispheres. Data are post-processed and binned to intervals of 15° in zenith angle and 30° in azimuth angle. In this study, data for the red (0.650–0.670 nm) and the near-infrared (810–840 nm) bands were used.

4) Soil data by Irons

Three data sets of a bare soil multiangular reflectance were collected by Irons et al. (1992) in 1989 using the MMR instrument on the ground and for several different sun zenith angles. The three data sets differ in the surface roughness of the soil, which was produced by working the soil with different agricultural machines. Surface roughnesses were 1.2, 2.6, and 3.9. The view zenith angle ranged from 0° to 70° in intervals of 10° and the view azimuth angle from 0° to 180° in intervals of 45°.

5) Grassland data from FIFE

The reflectance of a grassland at the FIFE study site located south of Manhattan in the Konza Prairie of northeastern Kansas was measured on July 11 and October 9, 1987 using MMR (cf. Walthall and Middleton, 1992). Vegetation was primarily a mixed grass including several species. View zenith angles observed were 0°, 20°, 35°, and 50°, the solar zenith angle was around 20° for one data set and around 55° for the other.

6) Airborne POLDER data

An airborne version of the POLDER instrument was used to collect multiangular reflectance data in the area of La Crau, France, in June, 1991 (Leroy et al., 1996). The area is covered by a wide variety of vegetation types, such as sorghum, sunflower, vegetable, vine and grass. After registration of several POLDER images, data with various view and illumi-

nation angles were obtained for the red (630–670 nm) and near-infrared (830–870 nm) wave bands.

EVALUATION OF MODEL PERFORMANCE

RMSEs of Inversion and Kernel Selection

To evaluate the four available kernel combinations for the Ambrals model, each was fitted to each of the multiangular data sets and the RMSEs and correlation coefficients between the predicted and observed reflectances were computed. Inversions were carried out simultaneously for all available bands, i.e., each kernel combination was required to not only fit the reflectances in each band well, but also to minimize the error across bands (mostly two, red and near-infrared) according to equation (4) at the same time. The results show that nearly all data sets are fit well by at least one of the model variants.

This may be seen in Figure 1, which shows the RMSE for each model variant for selected different land cover types. As expected, different kernel combinations produce the best fit (the lowest RMSE) for different types of land cover.

Cases with bare soil or sparse vegetation, such as the barren plowed field, the bare soil (Irons), the annual grass (coverage 4 percent), or the hard wheat (coverage 11 percent) are fitted better by the Ambrals variants with the Li-sparse kernel than those with the Li-dense kernel. For these land cover types, shadow-casting of clumps of soil and vegetation mainly affects the bidirectional reflectance. Thus surface scattering dominates the scattering of solar radiation. Due to the smallness of the contribution due to volume scattering, the selection of volume scattering kernel has little influence. Of the two surface scattering kernels, the Li-sparse kernel is chosen over the Li-dense kernel because it is the one that most strongly derives its shape from effects of shadow-casting.

For horizontally uniform vegetation canopies with many leaves, the type most different from bare soil and sparse vegetation, results are appropriately different. Examples shown in Figure 1 are the orchard grass (LAI is 1.0) and the irrigated wheat (LAI is 4.0). Their canopies tend to be continuous and thus volume scattering is dominant. Whether the Li-sparse or the Li-dense kernel is selected does not make much difference in these cases. But the model variants containing the Ross-thick kernel fit the bidirectional reflectances better than those using the Ross-thin kernel.

A third type of canopy is represented by dense vegetation composed of individual crowns, where mutual shadowing in viewing and illumination direction is the dominant process. In such cases, only illuminated tops of crowns are visible at large view zenith angles. Shadows cast by the crowns are mostly invisible due to mutual overlapping in the view direction, and no background is visible, which makes this case different from the case of sparse vegetation, where the shadows remain visible. As expected, model variants using the the Li-dense kernel are found to provide the best fit to the dense forest canopy, such as the hardwood forest data set.

But for the sparser old black spruce stand, shown in Figure 1 in contrast to the hardwood forest, both mutual shadowing of crowns and shadow casting play a role. Thus the model variants with the Li-sparse kernel fit this data set a little better than those with the Li-dense kernel, although all four fit it well.

For some land cover types, all four model variants fit at almost the same level. Examples are the soybeans (data from Kimes and Ranson), the grassland (FIFE) and the sunflowers (POLDER). This may be because the respective canopies do not display strong bidirectional reflectance properties for the given angular samplings.

From this analysis it is obvious that the BRDFs of different land cover types are best represented by different kernels in the Ambrals model. No data set is fitted with an RMSE of more than 0.046. The pattern given by which kernels fit well and which do not in a particular case can indicate basic characteristics of the observed surface, information that is potentially useful in land cover classification since it is different from the spectral information (for example the sparse/dense distinction for forests). What is retrieved is not merely the parameters required for a reasonable fit, but the kernel choice also reveals whether one particular approximation made or the other is more suitable for describing the BRDF. Such information would not be obtained in this way from a single BRDF model variant, especially since all Ambrals kernel combinations produce fits RMSEs in a seemingly reasonable range.

Principal Plane and Principal Cone Fits

This leads to an important question regarding the RMSEs of the four Ambrals kernel combinations plotted for each data set in Figure 1. Do the relatively moderate differences seen in the value of the RMSE between the four model variants signify relevant differences in the modeled reflectances? To answer this question, Figure 2 shows principal plane and principal cone plots of the modeled reflectance for the best-fitting and the worst-fitting kernel combinations, and the observed data, for a few cases and in the red and near-infrared wave bands. These cases were selected to have similar solar zenith angles of observation and to represent major types of vegetation, bare, sparse, dense and forest.

The bare soil and very sparse annual grass data sets were best represented by models containing the Li-sparse kernel. Figure 2 shows this fit and, for comparison, that obtained when using the Li-dense kernel instead. Obviously, employing the Li-sparse kernel provides a better fit in the hotspot region. For the annual grass, the azimuthal dependence of reflectance on the principal cone is not modeled too well by the model variants containing the Li-dense kernel.

The plots for the dense irrigated wheat show that the model variant based on the Ross-thick kernel fit the azimuthal change of reflectance on the principal cone much better than those containing the Ross-thin kernel, use of which leads to a clear underestimation of the reflectance.

Finally, the plots for the hardwood forest show that while both the best-fitting model variant, Ross-thick/Li-dense, and the worst-fitting model, Ross-thin/Li-sparse, have some problems with this data set, the fits provided by the former are better than those of the latter, especially on the principal plane in the near-infrared.

Figure 2 serves to demonstrate that not only do different types of land cover respond to different kernel combinations used in the Ambrals model, but the combinations producing the smaller RMSE also fit the data better to an extent that justifies, we believe, using different kernel combinations depending on the case. Even though the RMSEs of the bad-fitting kernel combinations are not much higher than those of the best-fitting combinations, the fits produced are not of similar quality. Since cases may be found where one or the other kernel combination works best, all should be retained in modeling.

Consequently, the best-fitting model variant in terms of the RMSE is chosen for the subsequent analysis. Table 1 identifies the best kernel combination and the band-averaged RMSE for each data set found from the inversion. In two cases, the best fit was obtained when the observations were modeled purely from volume scattering.

Correlation Between Modeled and Observed Data

For each data set the correlation coefficient between modeled and observed reflectances is calculated in the red and the near-infrared. The results are also shown in Table 1. In the red band, nearly half of all data sets, 12 of 27, had correlation coefficients larger than 0.9; 18 of 27 had coefficients larger than 0.8; and 23 of 27 sets had coefficients larger than 0.75. Only one set had a correlation coefficient less than 0.7, which will be explained later. In the near-infrared, the fits are even better. Half of all sets, 13 of 27, have correlation coefficients larger than 0.9. 23 of 27 have coefficients larger than 0.8, and 26 of 27 have coefficients larger than 0.75. These values indicate a reasonable agreement between the modeled and observed values.

Scatter plots of modeled versus observed reflectance are shown in Figure 3. For each of four different land cover types, bare, forest, broadleaf crops, and grasses, the data with the best correlation coefficient and the data with the worst correlation coefficient are shown to demonstrate the range of results obtained. The bare soil is modeled quite well in both cases, with correlation coefficients over 0.9. For the other land cover types, the good fits are excellent while for the ones with lower correlation coefficients display some scatter of points away from the diagonal. In the case of the pine forest, it is probable that the insufficient handling of multiple scattering by the kernels, which approximate it as being isotropic, causes the problems. An isotropic approximation to multiple scattering may be reasonable for sparse canopy cases, but the pine forest has a coverage of 70% and the data were obtained at four sun zenith angles ranging from 26° to 74°. The effect of multiple scattering may be expected to be large. We will later show that an im-

provement of the multiple scattering treatment of the model will indeed improve this fit.

In the broadleaf crop class, the corn data produces the worst fit, with a correlation coefficient in the red of 0.47. However, these data are extremely noisy, either due to problems during the measurement or due to actual local peculiarities of the canopy viewed that are probably not relevant on the larger scale of satellite remote sensing. Figure 4 plots a part of these data in the red to illustrate the point; the results for this data cannot be taken as typical in any way, or as testing a model. The unsystematic change of reflectances observed in the red band is not present as strongly in the near-infrared band, which promptly produces a better fit, the correlation coefficient being 0.75. Roujean et al. (1992), testing their model on this data set, also found correlation coefficients of the same magnitude, 0.37 and 0.73. Consequently, instead of this data set, Figure 3 displays the next-worst fit for a broadleaf crop, that for soybean data (Kimes).

A similar situation is encountered with the worst example of a fit to grass-like vegetation, where the grass observed by POLDER has a correlation coefficient in the near-infrared that is singularly small. It is possible that registration errors cause noisiness in the data in this case. Here, as in the corn data case, the fit is greatly improved when in inversion different weights are given to individual observations, these weights having been chosen according to the distance each data point has from the local mean. The next-worst typical case, shown in Figure 3, is the lawn data set (Kimes).

The overall conclusion is that, with the exception of the two data sets mentioned that have other problems and should be treated with caution, all land cover types investigated may be reasonably well represented by the Ambrals BRDF model. From Figure 2 one may note that different shapes of the BRDF are represented, such as ones with a hotspot (field, annual grass) or a bowl shape (irrigated wheat, hardwood forest).

This paper is concerned only with the Ambrals model, but it is interesting to see how it compares with one other semiempirical model, the three-parameter RPV BRDF model developed by Rahman et al. (1993) and modified by Martonchik (Engelsen et al., 1996). The RMSEs and correlation coefficients given in Table 1 cannot be directly compared with values published for the RPV model, since the orginal publication (Rahman et al., 1993) is not based on the modified version, whereas the publication using the modified version (Engelsen et al., 1996) postprocessed the data sets used to correct for diffuse skylight, a method not employed in this study to avoid using a standard atmosphere for data sets collected under various conditions. Therefore, table 2 gives RMSEs and correlation coefficients for the datasets measured by Kimes and co-workers for the Ambrals and the RPV model in direct comparison. In terms of the band-averaged RMSE, the Ambrals model has a lower RMSE in 9 out of 11 cases, the other two being ties. The average RMSE from the Ambrals model is 0.034, that from the modified RPV model 0.041, or 20 percent more. In terms of the correlation coefficient, in the red band, the Ambrals model produces better correlation between measured and observed reflectances for 8 of 11 data sets, with one tie. In the near-infrared band, Ambrals produces better fits in 9 of 11 cases, also with one tie. Judged by this particular series of data sets, the Ambrals model seems to be somewhat better. In many cases, however, the differences are not large. The modified RPV model clearly is similarly capable of also generating good fits to the data.

IMPROVED TREATMENT OF MULTIPLE SCATTERING

In the kernel-driven semiempirical modeling approach, multiple scattering is generally assumed to be isotropic, covered by the isotropic constant of the model. In reality, however, multiple scattering is dependent on the sun zenith angle. In some applications, this dependence is weak enough to be ignored, but in others it is not. When fitting the models to data covering several different sun zenith angles, problems may occur.

We study this problem further by using a fourth kernel term in the models to describe the sun zenith angle dependence of multiple scattering more explicitly. This kernel, the Hapke kernel, is derived from theory put forward by Hapke (1981). Based on the fundamental principles of radiative transfer theory, Hapke derived an analytical equation for the bidirectional reflectance function of a semi-infinite medium.

The single-scattered radiance is derived exactly. A two-stream approximation is used to calculate the multiply scattered radiance of isotropic scatterers. From the resulting expression, a kernel is derived that has the following form:

$$R_H = \frac{1 - \sqrt{1-\omega}}{1 + 2\cos(\theta_i)\sqrt{1-\omega}}, \quad (5)$$

where ω is the single-scattering albedo.

Since ω cannot be a free parameter if this term is to be used in a kernel-driven model adhering to the form given by equation (2), it is set to fixed values in each wave band, here chosen to be 0.08 in the red and 0.8 in the near-infrared.

This term is dissimilar to the other kernels and may therefore be used in a 4-parameter BRDF model. Figure 5 compares inversion results from the 3-parameter and the corresponding 4-parameter model for several different data sets using scatter plots. For sparse vegetation, here represented by hard wheat (coverage is only 11 percent) and steppe grass (coverage 5 percent), no difference is observed between modeling with and without the Hapke kernel. This is expected, since multiple scattering plays only a negligible role for sparse vegetation. For dense vegetation, however, the accuracy of the fits increases. Coefficients between modeled and observed reflectances increase in both the red and the near-infrared bands, for example from 0.775 to 0.880 in the red and from 0.715 to 0.804 in the near-infrared for the pine forest data set. For the lawn grass, the increase is from 0.724 to 0.787 in the red, and from 0.835 to 0.955 in the near-infrared. Similar results apply to the soybean data set. These results demonstrate that in cases of dense vegetation, adding one parameter related to multiple scattering can improve the fit quality of the model, whereas no change is achieved for sparse vegetation.

Caution should be used when applying the Hapke kernel to reflectance data remotely sensed from space. The Hapke kernel depends on a determination of the solar zenith angle dependence of the reflectance. However, in many remote sensing applications no range or just a rather small range of solar zenith angles is available in a particular period of time, for example for AVHRR, MODIS or MISR. As a consequence, the weight of this kernel will be ill determined for such applications. A 3-parameter model will allow more stable retrievals.

POSSIBLE IMPROVEMENTS TO THE INVERSION PROCESS

This paper has shown that all 27 data sets used were fitted well by at least one of the four kernel combinations that comprise the Ambrals BRDF model. This model may be used reliably for correcting angular effects in remote sensing applications. The capability of the model to describe bidirectional reflectance properties of surface objects can be further used to infer basic properties of the land covers viewed.

However, the inversions could perhaps be further improved in future work by considering the following factors. Firstly, the data used are field-measured. Radiation reaching the earth surface includes both direct and diffuse radiation owing to atmospheric scattering, which results in a partial smoothing of bidirectional reflectance. The data could be corrected for this effect where atmospheric characterization is available for the respective times and locations when the measurements were made. Using a standard atmospheric model could be considered in the absence of such data (Engelsen et al., 1996).

Secondly, the error function used in model inversion has an important influence on the resulting fits and the choice of kernels used. The weights W_j (see equation (3) qualify the contribution of every measurement to the error function. This may be used to improve the inversion, for example when a data set with nonuniform angular sampling displays a cluster of observations in a particular angle range. Most of the clustered observations do not contribute to establishing the shape of the BRDF but affect the error function unless they are given a smaller weight. Similarly, if a particular observation out of a set is considered to be noisy, the information it contributes to the inversion should be lessened by giving it a small weight. One possible way to approach this problem is to select a suitable weight for each observation based on an initial regression analysis. That is, the model is first inverted using the same weight for all observations, then weights are attributed to each observation according to a statistical criterion that identifies outliers, then the inversion is repeated. A different approach is to use error functions that are inherently more stable against noise and outliers (Tarantalo, 1987).

The problem of how to weigh the RMSEs achieved

in individual bands for deriving a band-averaged RMSE on which to base kernel selection has already been mentioned. Whether the difference in the magnitude of reflectances in the visible and in the infrared wavelength regions should enter the error function in form of a weighting term depends on whether a relative deviation (implying larger absolute errors in the infrared) or an absolute error (implying larger relative errors in the red) are preferable. Large relative errors are less favorable for deriving albedo from the BRDF, large absolute errors for deriving small reflectance values. In the present study, this problem is also visible in that the kernel combination best fitting both bands simultaneously is the combination fitting the near-infrared band for all 11 data sets by Kimes and co-workers, whereas it is also the best-fitting combination for the red band in only 7 of the 11 cases (and second-best fitting in 3 other cases). Thus for at least some types of vegetation, a suitable weight value should be added to the individual bands to avoid biasing the kernel selection towards the near-infrared reflectances.

Finally, the Li-kernels used contain two kernel-internal parameters describing crown shape and relative height that were each set to fixed values for sparse and dense canopies based on general considerations. If prior structural knowledge for specific land cover types is available, these parameters could be set to more appropriate values than the current ones. Wanner et al. (1995a) show that the crown shape in particular has an influence on BRDF shape.

MODELING BRDF EFFECTS IN AN ASAS IMAGE: A DEMONSTRATION

We demonstrate operation of the Ambrals model on a remotely sensed data set acquired by the airborne Advanced Solid-State Array Spectroradiometer (ASAS) (Irons et al., 1991) over the Walnut Gulch area in Arizona. Multiangular observations were acquired during overflights at a sun zenith angle of 38° and in the solar principal plane. The ASAS instrument was set up to acquire seven discrete images of the area at zenith angles 15°, 30° and 45° both forward-looking and aftward-looking, and at nadir. The ground spatial resolution at nadir was 2.01 m along the flight direction and 4.25 m across. All images were registered to the nadir image, and Ambrals model BRDF inversions were carried out on each pixel.

Figure 6a shows a spatial mosaic composited of three bands with center wavelengths of 549, 661, and 787 nm. The right half of the image is the data acquired at 45° zenith angle in the backscattering direction, where the sun is behind the sensor. The left half of the image is data acquired at 45° forward scattering zenith angle, where the sensor is facing the sun. For obvious reasons, more shadows are visible in the image showing the forward-scattering reflectances (left half) than in the image showing the backscattering reflectances (right half). A clearly visible seam runs through the image where the data from the two different viewing directions meet. This seam illustrates the magnitude of the BRDF effect present in a surface such as this one if not removed. Clearly, angular effects would have an impact on land cover classification and image interpretation if they were not corrected. Seams like this are known from AVHRR data in areas where data from different orbits, implying different viewing geometries, are mosaicked together (see Li et al. (1996) for an example and discussion), and will also occur for the similarly cross-track scanning MODIS instrument.

After fitting the Ambrals model to the string of 7 bidirectional reflectances available for each pixel, the model was used in forward mode to predict the reflectances for the left half of the image for the backscattering view zenith of 45°, corresponding to the angle at which the right half of the image was acquired. If the Ambrals model fails to correctly model the BRDF seen, the predicted reflectances and the resulting image will be noisy or the seam will be still visible. However, Figure 6b shows the composite, where the left half is the forward-modeled data and the right half is the same data as shown in Figure 6a. The seam is gone. Shadows vanish where they should as the model takes the reflectances into the hotspot. Instead of shaded backsides, illuminated front sides of objects and slopes are now visible. This demonstrates once more that the Ambrals model is indeed capable of modeling correctly the reflectances observed.

Figure 6c shows the nadir reflectance predicted for each pixel by the Ambrals model where the inversion is carried out using the 6 bidirectional reflectances excluding the nadir observation. Figure 6d shows for comparison the actually observed nadir image. The two are clearly very similar, illustrating that nadir

reflectance can be predicted from a string of observations at different angles similar to angles applicable to AVHRR and MODIS observations. Again, a failure of the Ambrals model to produce a reasonable fit for each pixel would have resulted in a noisy image.

CONCLUSIONS

In this paper, we have analyzed the capability of the kernel-driven semiempirical Ambrals BRDF model to provide adequate mathematical descriptions of the anisotropic reflectance of a variety of natural surfaces. Kernel-driven models combine advantages of physical models and empirical models in that they are highly adaptable to a large variety of occurring BRDF shapes, especially to BRDFs of different land cover types and of mixed pixels, but retain a basically-physical interpretation of the shapes produced. They can be inverted analytically through matrix inversion, they scale spatially and possess only three parameters, which is probably the maximal number that can reliably be inverted from the limited angular sampling available from most space-based instruments. The Ambrals model will thus be used in producing the global MODIS BRDF/albedo standard data product.

The mathematical expressions used in the Ambrals model introduced by Wanner et al. (1995a) are validated in this paper using 27 different measured BRDF data sets of a large variety of land covers. We find that the Ambrals model is fully capable of modeling these BRDFs with reasonable accuracy, the RMSEs being 0.034 on the average and correlation coefficients between modeled and observed data being larger than 0.8 or even 0.9 in a great majority of cases. Furthermore, the kernels selected may tentatively be related to vegetation structural characteristics, with differences observed between shadow-casting canopies and those where mutual obscuring of objects occurs, and between shadow-casting and strongly volume-scattering canopies. While several avenues for further improving the inversions have been discussed, the models in their current form are well-developed enough to be applied to remote sensing problems involving the extraction of the BRDF and the correction of multiangular imagery.

We would like to thank all investigators who allowed us to use BRDF data they collected, especially Jim Irons for making available the ASAS data of Walnut Gulch. We also thank the members of the MODIS BRDF/albedo team for their support and numerous discussions. This work was supported by NASA under NAS5-31369.

REFERENCES

Deering, D. W., S. P. Ahmad, T. F. Eck, and B. P. Banerjee, Temporal attributes of the bidirectional reflectance for three boreal forest canopies, *Proc. Int. Geosci. Remote Sens. Symp. 95*, 1239–1241, 1995.

Deschamps, P. Y., F. M. Breon, M. Leroy, A. Podaire, A. Bricaud, J. C. Buriez, and G. Seze, The POLDER mission: Instrument characteristics and scientific objectives, *IEEE Trans. Geosci. Remote Sens., 32*, 598–615, 1994.

Diner, D., C. J. Bruegge, J. V. Martonchik, G. W. Bothwell, E. D. Danielson, V. G. Ford, L. E. Hovland, K. L. Jones, and M. L. White, A multi-angle image spectroradiometer for terrestrial remote sensing with the Earth Observing System, *Int. J. Imag. Sys. Tech., 3*, 92–107, 1991.

Engelsen, O., B. Pinty, M. M. Verstraete, and J. V. Martonchik, Parametric bidirectional reflectance factor models: evaluation, improvements and applications, *Report, Joint Research Centre of the European Commission, EU 16426,*, 114 pp., 1996.

Hapke, B., Bidirectional Reflectance Spectroscopy, 1. Theory, *J. Geophys. Res., 86*, 3039–3054, 1981.

Hu, B., W. Wanner, X. Li, and A. Strahler, Validation of kernel-driven semiempirical BRDF models for application to MODIS-MISR data, *Proc. Int. Geosci. Remote Sens. Symp. 96*, 1669–1671, 1996.

Irons, J. R., K. J. Ranson, D. L. Williams, R. R. Irish, and F. G. Huegel, An off-nadir pointing imaging spectroradiometer for terrestrial ecosystem studies, *IEEE Trans. Geosci. Remote Sens., 29*, 66–74, 1991.

Irons, J. R., G. S. Campbell, J. M. Norman, D. W. Graham, and W. M. Kovalick, Prediction and measurement of soil bidirectional reflectance. *IEEE Trans. Geosci. Remote Sens., 30*, 249–260, 1992.

Kimes, D. S., Dynamics of directional reflectance factor distribution for vegetation canopies, *Appl. Opt., 22(9)*, 1364–1372, 1983.

Kimes, D. S., W. W. Newcomb, C. J. Tucker, I. S. Zonneveldt, W. van Wijngaarden, J. de Leeuw, and G. F. Epema, Directional reflectance factor distributions for cover types of Northern Africa, *Remote Sens. Env., 18*, 1–19, 1985.

Kimes, D. S., W. W. Newcomb, R. F. Nelson, and J. B. Schutt, Directional reflectance distributions of a hardwood and a pine forest canopy, *IEEE Trans. Geosci. Remote Sens., 24*, 281–293, 1986.

Leroy, M., and J.-L. Roujean, Sun and view angle corrections on reflectances derived from NOAA/AVHRR data, *IEEE Trans. Geosci. Remote Sens., 32*, 684–697, 1994.

Leroy, M., Surface reflectance angular signatures from airborne POLDER data, *Remote Sens. Environ., 57*, 97–107, 1996.

Leroy, M., J. L. Deuze, F. M. Breon, O. Hautecoeur, M. Herman, J. C. Buriez, D. Tanre, S. Bouffies, P. Chazette, and J.-L. Roujean, Retrieval of atmospheric properties and surface bidirectional reflectances over land from POLDER/ADEOS, *J. Geophys. Res., this issue*, 1997.

Lewis, P., The utility of kernel-driven BRDF models in global BRDF and albedo studies, *Proc. Int. Geosci. Remote Sens. Symp. 95*, 1186–1187, 1995.

Li, X., and A. H. Strahler, Geometric-optical bidirectional reflectance modeling of the discrete crown vegetation canopy: effect of crown shape and mutual shadowing, *IEEE Trans. Geosci. Remote Sens., 30*, 276–292, 1992.

Li, Z., J. Cihlar, X. Zheng, L. Moreau, and H. Ly, Detection and correction of the bidirectional effects in AVHRR measurements over northern regions, *IEEE Trans. Geosci. Remote Sens.*, in press, 1996.

Nilson, T., and A. Kuusk, A reflectance model for the homogeneous plant canopy and its inversion, *Remote Sens. Env., 27*, 157–167, 1989.

Rahman, H., B. Pinty, and M. M. Verstraete, Coupled surface-atmosphere reflectance (CSAR) model, 2, Semiempirical surface model usable with NOAA advanced very high resolution radiometer data, *J. Geophys. Res., 98*, 20,791–20,801, 1993.

Ranson, K. J., L. L. Biehl, and M. E. Bauter, Variation in spectral response of soybeans with respect to illumination, view and canopy geometry, *Int. J. Remote Sens., 6*, 1827–1842, 1985.

Ross, J. K., *The Radiation Regime and Architecture of Plant Stands*, 392 pp., Dr. W. Junk Publishers, The Hague, 1981.

Roujean, J. L., M. Leroy, and P. Y. Deschamps, A bidirectional reflectance model of the earth's surface for the correction of remote sensing data, *J. Geophys. Res., 97*, 20,455–20,468, 1992.

Ruiz de Lope, E. V., and P. Lewis, Monitoring land surface dynamics in the HAPEX Sahel area using kernel-driven BRDF models and AVHRR data, to be submitted to Remote Sens. Environ., 1997.

Running, S. W., C. O. Justice, V. Salomonson, D. Hall, J. Barker, Y. J. Kaufman, A. H. Strahler, A. R. Huete, J.-P. Muller, V. Vanderbilt, Z. M. Wan, P. Teillet, and D. Carneggie, Terrestrial remote sensing science and algorithms planned for EOS/MODIS, *Int. J. Remote Sens., 15*, 3587–3620, 1994.

Strahler, A. H., M. J. Barnsley, R. d'Entremont, B. Hu, P. Lewis, X. Li, J.-P. Muller, C. B. Schaaf, W. Wanner, and B. Zhang, MODIS BRDF/albedo product: Algorithm theoretical basis document, *NASA EOS-MODIS Doc. and update*, version 3.2, 65 pp., 1995.

Strahler, A. H., C. B. Schaaf, J.-P. Muller, W. Wanner, M. J. Barnsley, R. d'Entremont, B. Hu, P. Lewis, X. Li, and E. V. Ruiz de Lope, MODIS

BRDF/albedo product: Algorithm theoretical basis document, *NASA EOS-MODIS Doc.*, version 4.0, 1996.

Tarantola, A., *Inverse problem theory : methods for data fitting and model parameter estimation*, Elsevier Science Pub. Co., 1987.

Walthall, C. L., J. M. Norman, J. M. Welles, G. Campbell, and B. L. Blad, Simple equation to approximate the bidirectional reflectance from vegetation canopies and bare soil surfaces, *Appl. Opt., 24*, 383–387, 1985.

Walthall, C. L., E. M. and Middleton, Assessing spatial and seasonal variations in grasslands with spectral reflectances from a helicopter platform, *J. Geophys. Res., 97*, 18905–18912, 1992.

Wanner, W., X. Li, and A. H. Strahler, On the derivation of kernels for kernel-driven models of bidirectional reflectance, *J. Geophys. Res., 100*, 21077–21090, 1995a.

Wanner, W., A. H. Strahler, J.-P. Muller, M. Barnsley, P. Lewis, X. Li, and C. L. Barker Schaaf, Global mapping of bidirectional reflectance and albedo for the EOS MODIS project: the algorithm and the product, *Proc. Int. Geosci. Remote Sens. Symp. 95*, 525–529, 1995b.

Wanner, W., A. H. Strahler, B. Hu, X. Li, C. L. Barker Schaaf, P. Lewis, J.-P. Muller, and M. J. Barnsley, Global retrieval of bidirectional reflectance and albedo over land from EOS MODIS and MISR data: theory and algorithm, *J. Geophys. Res.*, in press, 1997.

Table 1: Summary of Data Sets, Kernel Selection, Inversion RMSEs and Correlation Coefficient (r) Between Modeled and Observed Reflectances

Data Source	Cover Type	Cov. %	LAI	θ_i, Range	Best-Fitting Kernel	RMSE	r (red)	r (NI
Kimes (1983),	Plowed Field	0	0	26-45	Ross-thick/Li-sparse	0.016	0.976	0.976
Kimes et al.	Corn	25	0.65	26-68	Ross-thin/Li-dense	0.028	0.467	0.750
(1985, 1986)	Lawn Grass	97	9.9	42-70	Ross-thin/Li-dense	0.046	0.724	0.835
	Soybeans	90	4.6	28-76	Ross-thick/Li-sparse	0.043	0.783	0.806
	Hardwheat	11		27-51	Ross-thick/Li-sparse	0.019	0.963	0.938
	Annual Grass	4		28-50	Ross-thick/Li-sparse	0.023	0.945	0.878
	Steppe Grass	5		35-63	Ross-thick/Li-sparse	0.024	0.887	0.922
	Irrigated Wheat	70	4.0	26-59	Ross-thick	0.037	0.911	0.917
	Orchard Grass	50	1.0	45-82	Ross-thick/Li-sparse	0.030	0.837	0.914
	Pine Forest	70		26-74	Ross-thin/Li-dense	0.041	0.775	0.715
	Hardwood Forest	79		25-79	Ross-thick/Li-dense	0.030	0.902	0.890
Ranson et al.	Soybeans	72	3.0	20-49	Ross-thick/Li-sparse	0.017	0.753	0.892
(1985)		83	3.9	21-38	Ross-thin/Li-dense	0.015	0.738	0.851
		99	2.9	31-61	Ross-thick	0.011	0.914	0.927
Deering	Aspen			45-59	Ross-thick/Li-sparse	0.025	0.922	0.883
et al.	Old Black Spruce			36-59	Ross-thick/Li-sparse	0.011	0.949	0.943
(1995)	Old Jack Pine			34-60	Ross-thin/Li-dense	0.010	0.899	0.947
Irons et al.	Soil I	0	0	16-68	Ross-thick/Li-sparse	0.027	0.920	0.915
(1992)	Soil II	0	0	34-53	Ross-thick/Li-sparse	0.016	0.971	0.970
	Soil III	0	0	28-54	Ross-thick/Li-sparse	0.027	0.931	0.933
FIFE (Walthall	Grass I		1.3	19-25	Ross-thin/Li-sparse	0.021	0.769	0.896
and Middleton, 1992)	Grass II		0.9	53-61	Ross-thin/Li-sparse	0.031	0.913	0.898
POLDER	Grass			37-47	Ross-thick/Li-sparse	0.042	0.706	0.560
	Sorghum			38-45	Ross-thin/Li-sparse	0.028	0.785	0.874
	Vineyard			38-45	Ross-thin/Li-sparse	0.024	0.809	0.934
	Vegetable			37-45	Ross-thin/Li-sparse	0.036	0.888	0.793
	Sunflower			38-47	Ross-thin/Li-sparse	0.023	0.869	0.910

Table 2: Comparison of Inversion RMSEs and Correlation Coefficients (r) Between Modeled and Observed Reflectances for the Ambrals Model and the Modified RPV Model

Data set	Ambrals			mod. RPV		
(Kimes)	RMSE	r (red)	r (NIR)	RMSE	r (red)	r (NIR)
Plowed Field	0.016	0.976	0.976	0.016	0.976	0.976
Corn	0.028	0.467	0.750	0.031	0.718	0.652
Lawn Grass	0.046	0.724	0.835	0.046	0.706	0.839
Soybeans	0.043	0.783	0.806	0.057	0.871	0.600
Hardwheat	0.019	0.963	0.938	0.025	0.941	0.897
Annual Grass	0.023	0.945	0.878	0.028	0.911	0.822
Steppe Grass	0.024	0.887	0.922	0.026	0.883	0.897
Irrigated Wheat	0.037	0.911	0.917	0.039	0.891	0.907
Orchard Grass	0.030	0.837	0.914	0.050	0.796	0.723
Pine Forest	0.041	0.775	0.715	0.049	0.659	0.560
Hardwood Forest	0.030	0.902	0.890	0.044	0.864	0.890

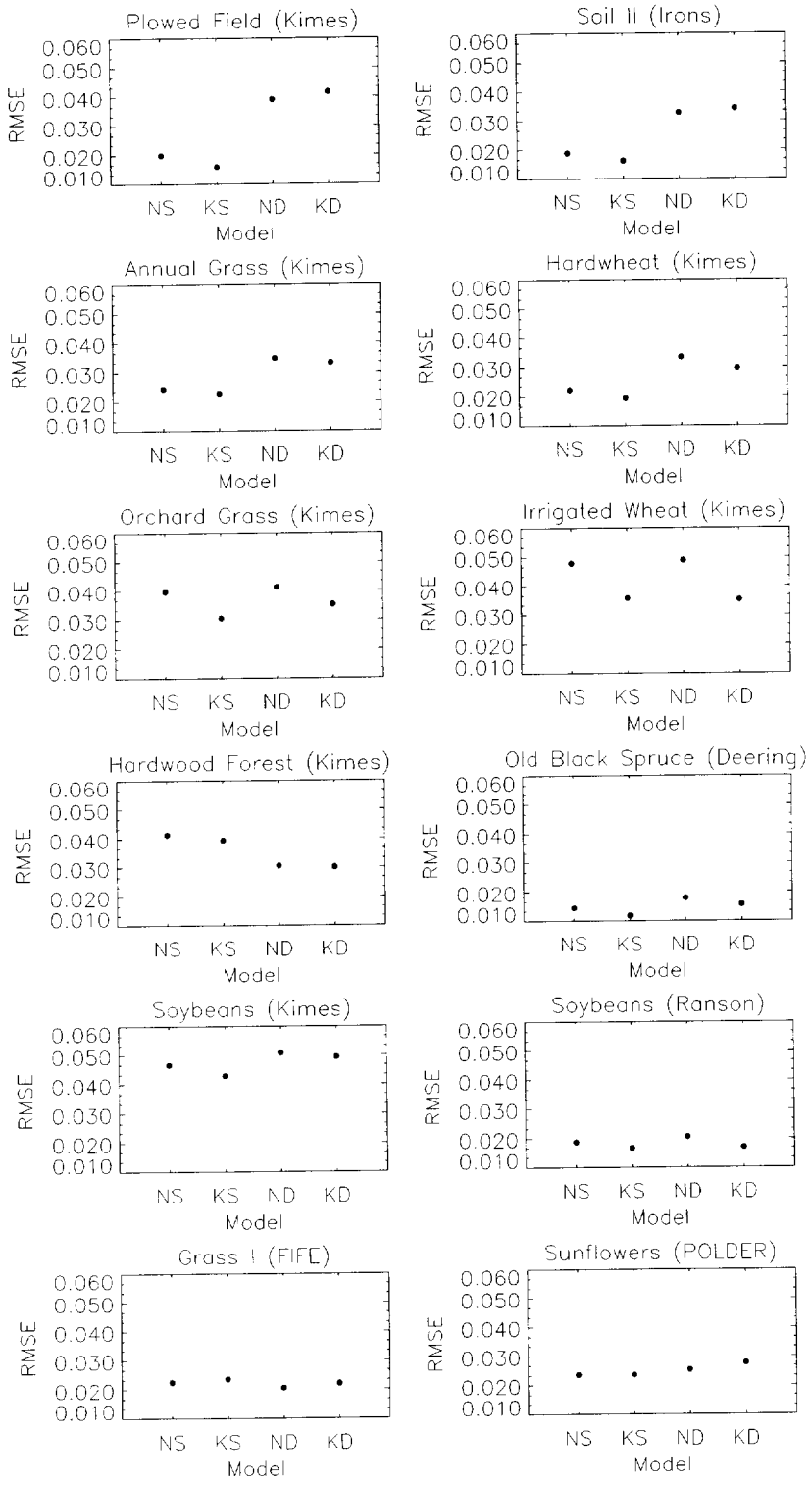

Figure 1: RMSEs of Ambrals model inversions for selected data sets representing different types of land covers (barren, sparsely vegetated, grass-like, forest, broadleaf crops; refer to Table 1 for coverages and LAIs). Kernel combinations are keyed as follows: NS, Ross-thin/Li-sparse; KS, Ross-thick/Li-sparse; ND, Ross-thin/Li-dense; KD, Ross-thick/Li-sparse.

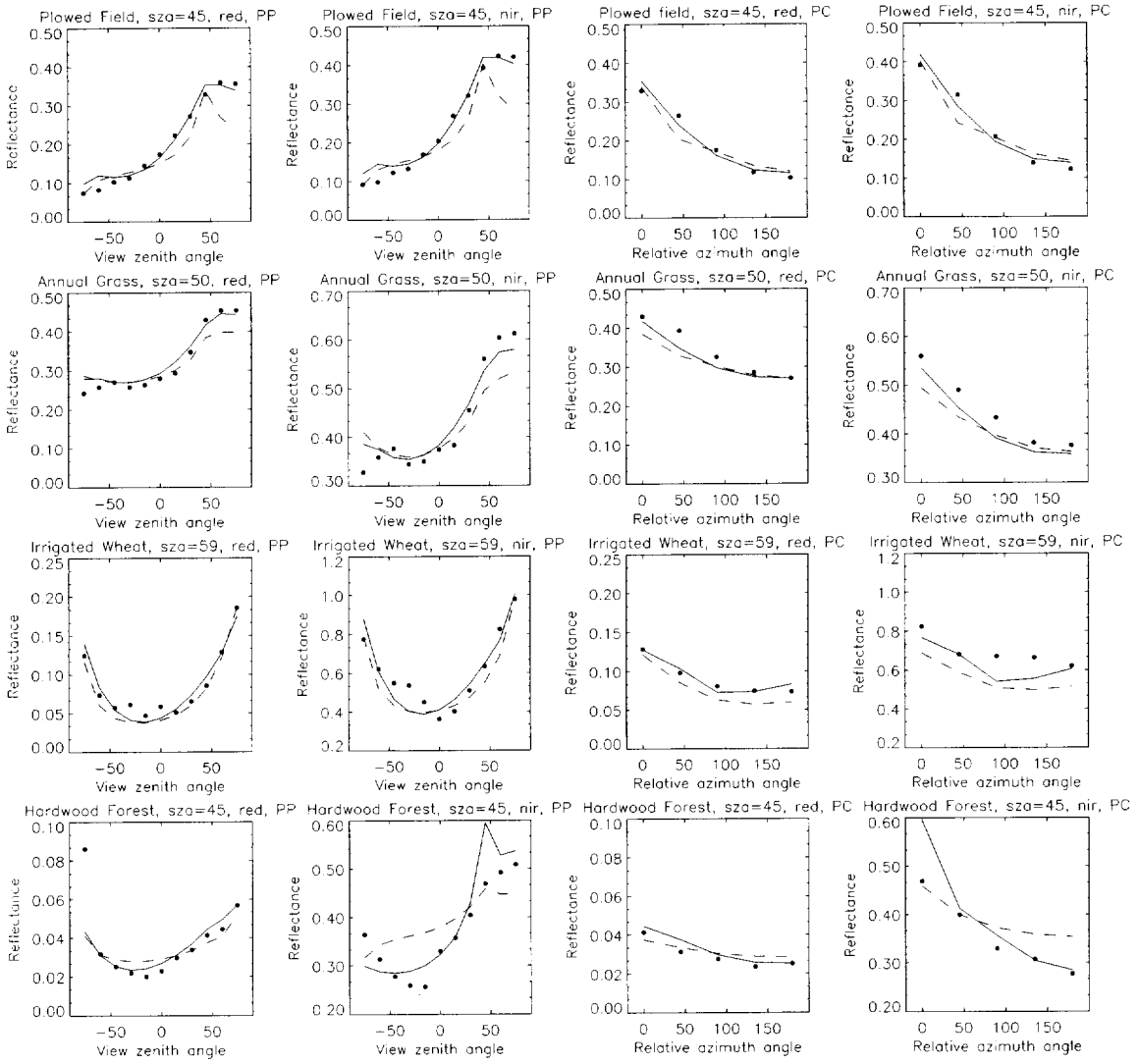

Figure 2: Comparisons of fits produced by good (solid lines) and bad (dashed lines) Ambrals kernel combinations on the principal plane and on the principal cone for selected data (dots) representing different types of land cover.

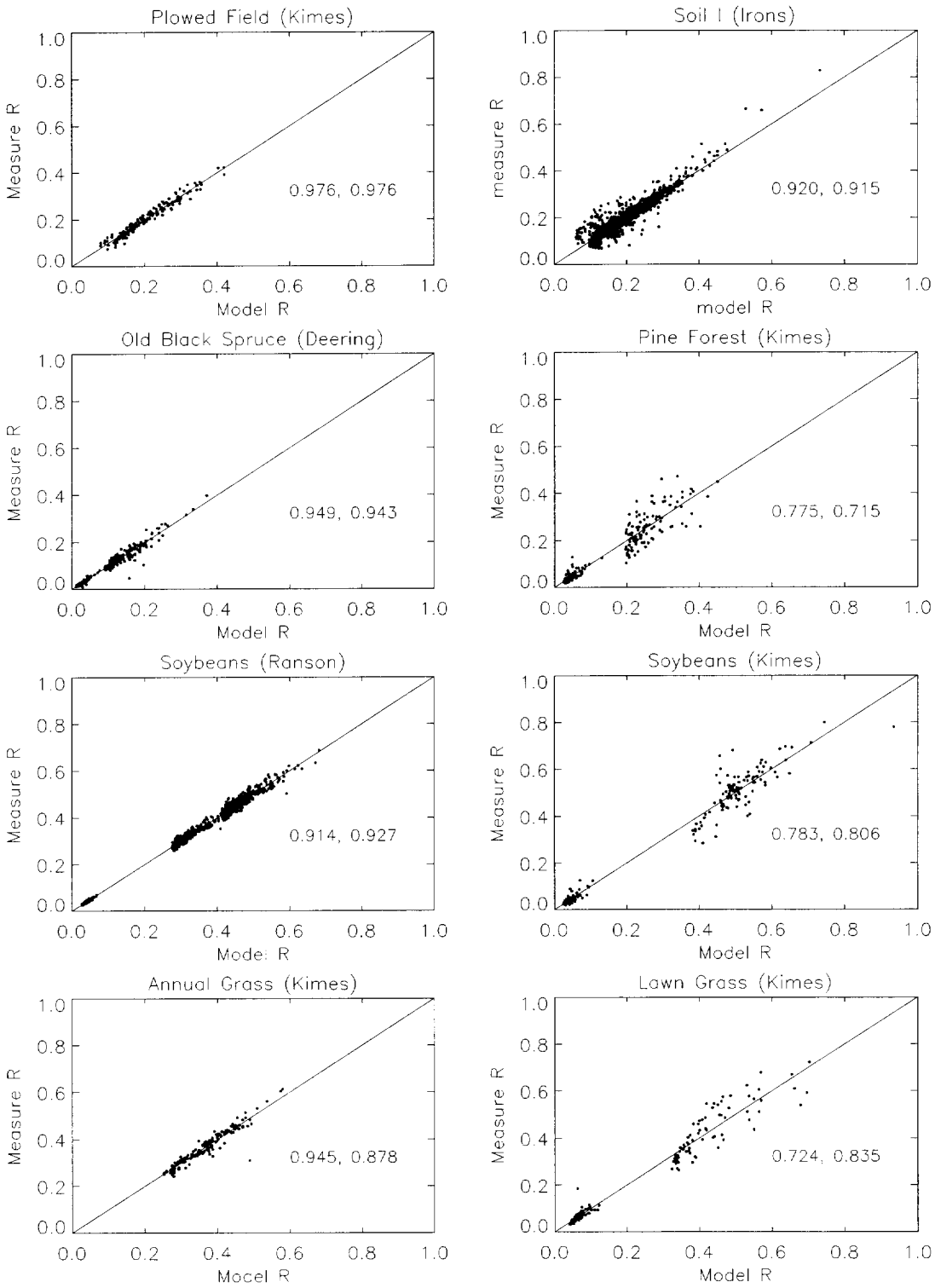

Figure 3: Modeled versus observed reflectances in the red and near-infrared for different types of land cover (barren, forest, broadleaf crops, grass-like). The numbers given in each panel are the correlation coefficient in the red and in the near-infrared bands, respectively.

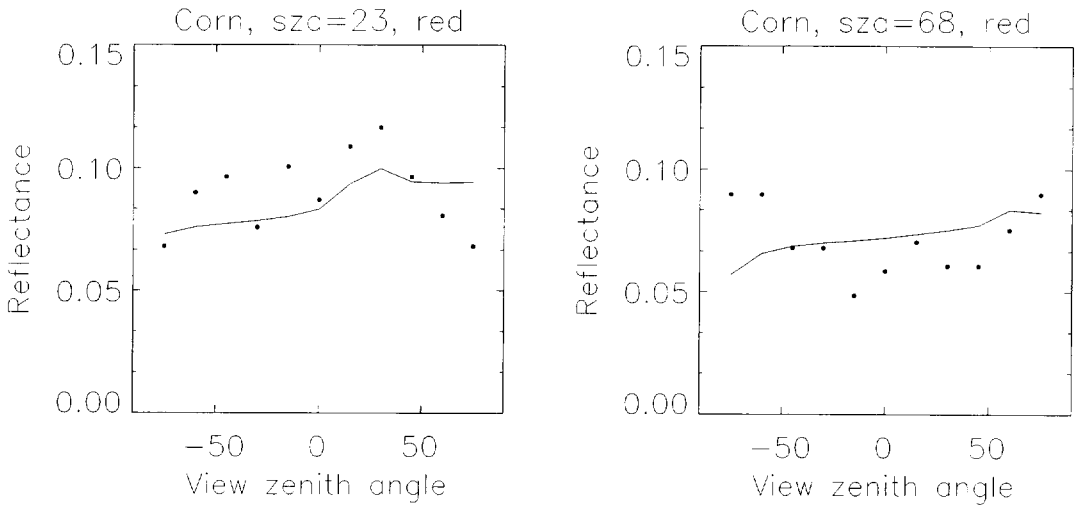

Figure 4: The corn data set shows irregular behavior of the observed reflectances (dots), leading to poor fits (solid lines). The situation is similar at other zenith angles in the red, and somewhat better in the near-infrared.

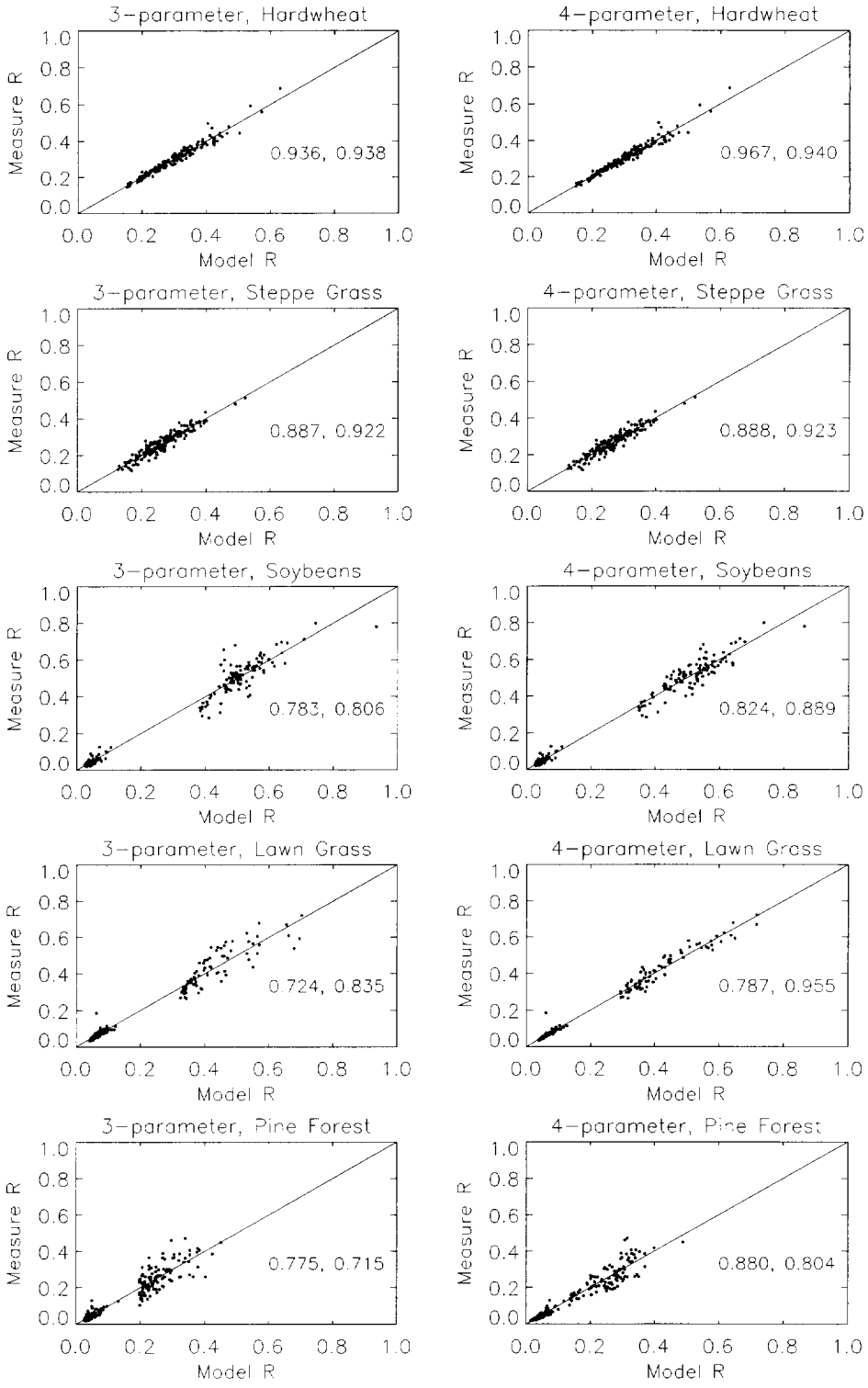

Figure 5: Modeled versus observed reflectances in the red and near-infrared for different types of land cover (top two rows: sparse vegetation; bottom three rows: dense vegetation) and for a 3-parameter model (left) and a 4-parameter model (right) where the Hapke-kernel for multiple scattering was added. The numbers given in each panel are the correlation coefficient in the red and in the near-infrared bands, respectively. Note that fits improve for the dense vegetation.

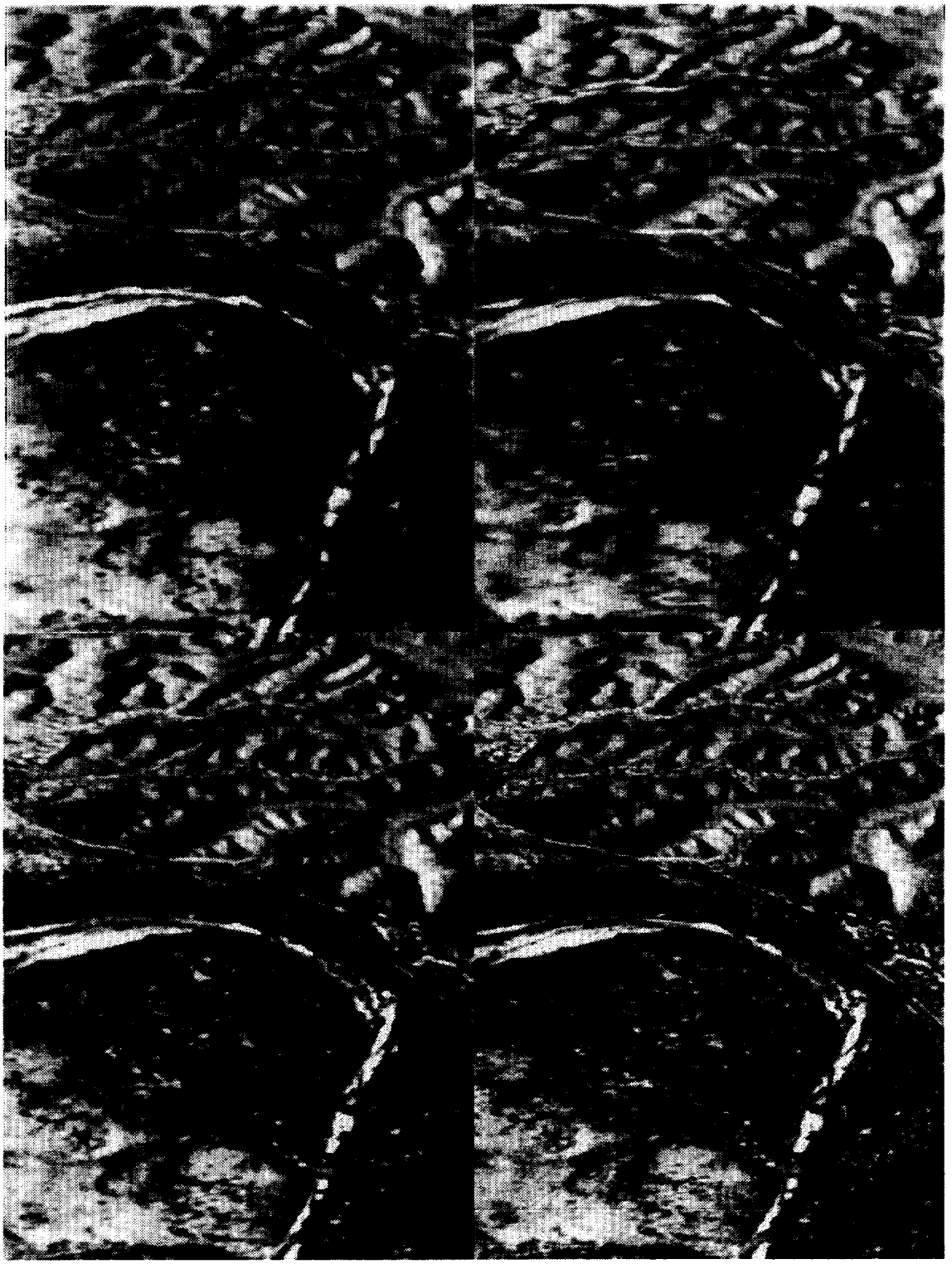

Figure 6: Grey-scale representation of a three-band composite image of ASAS data (wavebands centered at 549, 661, and 787 nm) over Walnut Gulch, Arizona. Sun zenith angle during data acquisition was 38°. Panel (a) (top left) shows in the right half data acquired at 45° zenith angle in the backscattering direction and in the left half data acquired at 45° zenith angle in the forward-scattering direction, where more shadows are visible. The second panel, (b) (top right), shows in its left half the same data as shown in the left half of panel (a), and in the right half data predicted for a backscattering view zenith angle of 45° using the Ambrals model on each pixel and inverting the seven reflectances observed by ASAS at 45°, 30°, 15°, and nadir in the forward and backscattering directions; i.e., the right half of this image is observed data, the left